

# Morality, utilitarianism, and rights

RICHARD B. BRANDT

*Professor of Philosophy (Emeritus), University of Michigan*



**CAMBRIDGE**  
UNIVERSITY PRESS

Published by the Press Syndicate of the University of Cambridge  
The Pitt Building, Trumpington Street, Cambridge CB2 1RP  
40 West 20th Street, New York, NY 10011-4211, USA  
10 Stamford Road, Oakleigh, Melbourne 3166, Australia

© Cambridge University Press 1992

First published 1992

Reprinted 1995

Printed in the United States of America

Library of Congress Cataloging-in-Publication Data is available.

A catalogue record for this book is available from the British Library.

ISBN 0-521-41507-1 hardback

ISBN 0-521-42527-1 paperback

# Contents

<i>Sources</i>	<i>page vii</i>
1 Introductory comments	1
A METAETHICAL PRELIMINARIES	
2 Moral philosophy and the analysis of language	19
3 Rational desires	38
4 The explanation of moral language	57
5 Morality and its critics	73
6 Rationality, egoism, and morality	93
B NORMATIVE ETHICS: UTILITARIANISM	
7 Some merits of one form of rule – utilitarianism	111
8 Fairness to indirect optimific theories in ethics	137
9 Two concepts of utility	158
C UTILITARIANISM AND RIGHTS	
10 The concept of a moral right and its function	179
11 Utilitarianism and moral rights	196
D DETERMINISM, EXCUSES, AND CHARACTER	
12 A utilitarian theory of excuses	215
13 A motivational theory of excuses in the criminal law	235
14 Traits of character: A conceptual analysis	263
15 The structure of virtue	289
E IMPLICATIONS OF UTILITARIAN THEORY	
16 The morality and rationality of suicide	315
17 Utilitarianism and the rules of war	336

## CONTENTS

18	Public policy and life and death decisions regarding defective newborns	354
19	Utilitarianism and welfare legislation	370
	<i>Index</i>	389

## Introductory comments

The articles on ethics collected in this volume,<sup>1</sup> with a few exceptions, represent views to which I subscribe today. The collection does not contain anything I have written on other topics: epistemology, metaphysics, or philosophy of mind. Various earlier pieces on ethics have also been omitted partly for lack of space and partly because I no longer like their content. But some omissions I regret are “The Concepts of Obligation and Duty” (*Mind*, 1964), “The Significance of Differences of Ethical Opinion for Ethical Rationalism” (*Philosophy and Phenomenological Research*, 1944), “The Emotive Theory of Ethics” (*Philosophical Review*, 1950), and “W.K. Frankena and Ethics of Virtue” (*Monist*, 1981), but these do not mark significant steps in the development of my mature thought. I mention some other omitted pieces below. I have not included any selections from my books; to do so would have been an enormous job and the books are readily available.

I paid little attention to the problems of ethics during my student years, at Denison University and Cambridge University; the interest began with seminars offered at Yale by W.M. Urban, with whom I thoroughly disagreed. This disagreement partly reflected some fundamental beliefs I had acquired earlier. One was that reliance on “intuitions” is no reliable basis for claims to truth of moral or value statements. I am afraid I now include in this dismissal the current view that we may find moral truth by reliance on the coherence of, or “wide reflective equilibrium” of, a person’s “considered” moral/value beliefs. A second basic belief is the counterpart of this, that nonlogical truth generally is to be obtained only by the methods of the empirical sciences – reliance on observation and the “best inference” from this. But I was never a logical positivist; indeed, my first publication in philosophy (in the *Journal of Philosophy* for 1938) was a criticism of the views of Waismann, early Wittgenstein, Carnap, and others, because they banned concepts that seem necessary for atomic physics. Third, I was unconvinced by arguments, such as those by Moore and Ross, that

1 It was Brad Hooker’s idea to publish this collection, and I am indebted to him. I am also indebted to Cambridge University Press for undertaking the publication and for their typically superior production performance.

## INTRODUCTORY COMMENTS

ethical terms name simple objective properties. With these assumptions, where should one move in moral philosophy?

I was first led to espouse a kind of emotive theory (*Ethics*, 1941), to the effect that, just as a sense experience is the ground of a perceptual judgment, so the occurrence of an emotion determines basic moral judgments, like "That was a contemptible deed." This was somewhat like Westermarck. In support of this view I argued that such moral judgments vary with the development of major "sentiments," with physiological states such as major depressions, with states of satiety or grief, just as do emotions. In a following publication (*Ethics*, 1946) I suggested that some evaluative adjectives (for example, "contemptible," "admirable") could be defined as "is a fitting object" of the corresponding emotion (contempt, admiration) – the one I had earlier argued determines the evaluation. I conceded that questions might be raised about the meaning of "fitting."

In the early years of my teaching at Swarthmore College (beginning in 1937) I had the good fortune of close association with several distinguished psychologists including Wolfgang Koehler, Karl Duncker, and (slightly later) Solomon Asch, all of whom had a strong interest in ethics, thought that basic value/moral judgments are the same the world over, and rejected learning theories inconsistent with this. A desire to put the absolutism of these psychologists to an empirical test, and to get light on the sources of moral judgments (it having been suggested by the anthropologist Ralph Linton and others that methods of child rearing had much to do with the development of value norms), led me to wide reading in anthropology and eventually to fieldwork among the Hopi Indians, reported in *Hopi Ethics: A Theoretical Analysis* (published in 1954). This work revealed the complications in appraising the absolutism of these psychologists, since divergent ethical standards usually occur in widely varying factual circumstances and hence in the presence of diverse factual beliefs about the topic, which could be explanatory of different specific norms even if everyone actually shared basic moral principles. I argued, however, that much evidence supports the view that some ethical norms vary even when factual beliefs are the same, especially ethical appraisals of cruelty to animals. I also uncovered evidence to show the role of personal advantage to individuals in the development of social ethical norms.

Since I now held that ethical appraisals do not always vary with the factual beliefs of judges about the situation being judged, and that hence we must have a theory of the learning of morality/values which permits this, and since I held that an emotional datum determines (some) value judgments, should I not accept the "emotive theory" of ethics as it was formulated, say, by Charles Stevenson? I objected to some parts of his view, and criticized these in a symposium with him in 1950 (*Philosophical*

*Review*, 1950). If Stevenson had merely said that moral statements "express" attitudes in the sense that a person who heard an ethical statement would take it that the speaker had a corresponding attitude (and would be somewhat moved to conform to it, depending on the prestige of the speaker), there might not have been disagreement. But Stevenson made detailed claims about moral language. For one thing, he held that hearing a moral statement made by another tends to elicit a conforming attitude, not because the auditor takes it a certain attitude has been expressed (by a prestigious person), but because of a *conditioned* response to the power of moral words, acquired in the process of language learning. I doubted whether there is reason to believe this. Second, he claimed that ethical disagreement must be viewed not as cognitive but as disagreement in attitude, partly because he thought facts *may* not settle *every* ethical dispute (with which I have noted that I agreed). He also argued, however, that once you have a proposal about what the cognitive meaning is, certain particular possible facts become relevant to settling a dispute, and only these, whereas we know that there is an unlimited range of reasons that may be relevant to an ethical dispute. I replied that this argument does not undermine either nonnaturalist accounts, or subjective naturalist accounts (to which Stevenson's own view was very close). Furthermore, I thought that not every "reason" that actually moves a person's view in a debate (for example, one that points to some personal advantage in preaching a certain view) is a morally relevant reason, as Stevenson thought (this debate was developed further in 1957 in *The Language of Value*, edited by Ray Lepley). I thought it somewhat cynical of him to insist, as he sometimes did, that moral disputes be viewed essentially like disputes between an employer and a labor union, each seeking an advantage for himself. Thus I did not go along with all of Stevenson.

Shortly thereafter I published a proposal (*Mind*, 1952) for how best to construe "assertion" and "cognitive meaning." I there proposed that a moral term *M* is best taken to designate some empirical property *P* for a given individual at a time if and only if it is causally impossible, at that time, that a speaker both assert "*S* is *M*" with assent or conviction and doubt that *S* is *P*. Is there any property *P* which meets this condition for moral predicates? Suppose it were suggested, for *S* is "wrong," that the property is "would be disapproved by anyone [or perhaps just by me the speaker] if he were fully informed on the relevant facts, were impartial . . ." (and so on for any properties it would be plausible to include in an absolutist form of subjectivist analysis). I replied to various objections, in defense of such an analysis, defending the possibility of a cognitive account of ethical statements, such as that of Roderick Firth (another colleague at Swarthmore), or Hutcheson, Hume, and Smith before him.

## INTRODUCTORY COMMENTS

In 1955, however (*Philosophy and Phenomenological Research*), I criticized Firth's proposal, partly for its details but primarily for its absolutism, because of doubts about the uniformity of responses among qualified persons.

In my *Ethical Theory* (1959) I proposed a kind of truce with the emotive theory, suggesting that one might agree that the primary function of ethical terms is to express attitudes, but that (along lines adopted by some writers at the time) the use of these special terms *implies* or *claims* that the attitude expressed has certain properties that intelligent, fully informed persons, who had thought through the problem of general policies for the endorsement of moral attitudes, would require for the appraisal of actions. I thought such persons would require that these "conditions" include full information, impartiality, and so on.

This is where I was in 1959. There ensued a rather gradual transition to the view expounded in the Locke Lectures (1974), which appeared in expanded form in 1979, in *A Theory of the Good and the Right*.

## A. METAETHICAL PRELIMINARIES

The first paper in this collection, "Moral philosophy and the analysis of language" (1963), takes a major step away from the view that *normative* reflection must *start* with some theory about the *meaning* of moral terminology. Some philosophers have thought that "good" just means pleasant, and "evil" painful (Bentham), or that "right" means maximizes the good (G. E. Moore), or that "*X* is the right act in *C*" means "I hereby prescribe doing *X* in circumstances like *C* (wherever they may be), and prescribe this overridingly" (Hare); and they go on from there to decide what is good or right. Why not? Suppose your language contained only one evaluative word, "prudent." Would you be willing to base your moral philosophy on this? Suppose speakers of English use moral words in different senses, some using "wrong" to mean forbidden by God. Would we then want to develop corresponding different theories of normative ethics? Suppose you light on a given analysis (as Firth did) by observing what one does to determine one's answer to a normative question. You can then never criticize a person's normative reasoning by appeal to what his terms mean, because your conclusion about his meaning rests on observation of the form of his normative reasoning.

There must be something better than reliance on prior conclusions about the meaning of moral language. But what is there? Several writers had already suggested a somewhat different approach: Frankena, Rawls, possibly even Kant and Mill. I suggested that one possibility is to get clear about what our problems really are. For instance, it seems the



## INTRODUCTORY COMMENTS

problem we are trying to solve when we ask what is good is about what to *desire*. If so, then we can ask what we would desire if we were fully informed about everything (the nature of the target, the etiology of the desire) that would cause a change in our desires – and this means bringing logic and facts maximally to bear on a desire. Such an inquiry might remove our doubt about what to desire. What else might we want to know? When we ask about what is *right*, perhaps what we want to know is which kinds of actions would be forbidden, or called for, by a moral system a person with criticized desires would *want* “written in the consciences of men.” (Assume here, building further on my old belief about the role of emotions in moral experience, that for an individual to embody a “moral system” is for him to be motivated to avoid actions of certain types, to feel guilty if he performs a type of action he is motivated to avoid, and to disapprove others who do, and to think that these attitudes are justified in some appropriate sense.) Finding what kind of moral system a person with criticized desires would want for his society would be bringing facts and logic to bear on the wanted system. Showing this would surely be a recommendation of that moral system, although none of this would show that some moral/value statement is *true*, roughly in the sense of the empirical sciences, unless we elected to define “good” or “morally right” so that it would be, as well we might (as I explain in a paper below).

Would we be satisfied to identify the main problems of ethics in this way? These suggestions add up to a big picture of what morality and values are, and how they may be criticized. Do we really like this picture? Is there some more satisfying way to look at things? Do we want to raise questions in our old terminology, ones not wholly provided for by the new conceptual scheme? My vote was to go for the new proposal, and still is.

The second paper, Chapter 3, “Rational desires” (1970), written after a move to the University of Michigan, was intended to supplement the preceding one, by showing how full information might be utilized to determine what an informed person would desire *for itself*. It has been questioned whether knowledge can function to criticize *basic* desires. It is here argued that such a desire is in the clear if it survives repeated confrontation with facts about it or its target.

This thesis depends on a part of the psychology of motivation. (Here I was influenced by the psychologists J. W. Atkinson and Albert Bandura, among others.) There are some experiences we like (dislike) natively (nice smells, being burnt), also some we like (dislike) as a result of conditioning (scotch and water). Suppose, then, we have an image or idea like one of these experiences. By conditioning or stimulus generalization this *idea*

## INTRODUCTORY COMMENTS

will seem attractive or have a favorable affective tone. That is what it is to desire that of which we have the idea. Sometimes an event will be wanted or aversive because of a relation of this sort to untypical or abnormal cases of the thing desired or aversive. What will happen if we face it with reflection on experiences that are typical? Answer: The desire (aversion) will tend to be extinguished either because of nonreinforcement or through counterconditioning (overlaying the desire/aversion with a more powerful conflicting motive). Many conforming desires/aversions we have learned because of scorn of parents or peers, when the relevant experiences would support desires/aversions of an opposite kind, can be labeled "unauthentic" and be extinguished by vivid awareness of the facts. A somewhat similar phenomenon is misgeneralization: Acquiring a desire/aversion to something of kind *ABC* when the initial disliked experience was to the more specific whole *ABCDE*. In more interesting cases mere reflection on what it was that one liked/disliked about the original situation can change the relevant desire. There are other types of situations in which the same sort of effect will occur.

I thus argued that vivid representation of full information about a situation can often modify desires for things for themselves. I suggested this is a valid form of *cognitive* criticism that can recommend, but it is not *evidence* supporting the truth of a value statement short of some definitions that make it so.

The third paper, Chapter 4, "The explanation of moral language" (1985), is to a considerable extent a summary and defense of my book, *A Theory of the Good and the Right* (1979). It defends a conception of "a person's morality" much like that stated above, and proposes that a major job of moral philosophy is, not to attempt to analyze moral terms as now ordinarily used in commonsense, but to find a pattern of sharp concepts useful in view of what morality is: for discussion on whether someone's wrong conduct should be excused, for giving moral advice, for personal reflection on what is right. For this we need only one basic concept, "fitting," since, as A. C. Ewing urged, all other moral concepts can be defined by this concept plus empiricist language. Thus for *X* to be "reprehensible" for his act *A* is for it to be fitting that persons disapprove of *X* because of *A*. According to Ewing, "fitting" names an unobservable relation. It appears, however, that "fitting" can be explained in terms of the notion of a moral system being "justified." And we can explain "justified" so that the concept is clear, does not purport to name an unobservable property or its application require intuitions, and it is normatively neutral (in the sense, for example, that it is not explained as "maximizes the happiness of society"). What kind of explanation of this sort can be given? I proposed "A moral code is justified

## INTRODUCTORY COMMENTS

for *X*" if *X* would choose to support it for his society if he were fully rational (with "rational" needing explanation). I assumed that if a moral system is justified in this sense for someone, that fact will recommend it to him. This proposal does not imply that a moral system justified for someone is justified for everyone, and the above account is designed to permit construction in a relativistic way, since "*X*" does not mean "just anyone." I suggested that people who understand what morality is would be satisfied with construing moral questions in this way. This set of definitions, if widely accepted, would prevent confusion in moral thinking and would make clear that morality is a rational phenomenon.

The fourth paper, Chapter 5, "Morality and its critics" (1989), in the form of a criticism of the philosophical views of Professor Bernard Williams, assembles a number of views I have wished to support. I have already suggested above that a person's morality consists of intrinsic aversions to some types of actions and corresponding dispositions to feel guilty and to disapprove of others and to think these attitudes are justified in some way. Such a system can be quite complex, perhaps nearly as complex as language (although it has to be capable of being learned). This, I suggest, is what a person's *morality* is. When there is a conflict between some aversions, the system may tell what to do by providing a "holistic" response to an individual situation. There are other standards in society, like etiquette or custom, for breach of which one can be criticized and may react by feeling shame or embarrassment, but these are not thought to be *justified* and are not part of the moral system.

Some philosophers have thought that when morality contains an aversion, for example, to harming others, it can be right to act contrary to it only when there is some stronger contrary aversion. (W. D. Ross says there is a real duty to do that, the doing of which is a *prima facie* duty, unless there is a stronger contrary *prima facie* duty.) Williams rightly denies this, but since this is not what morality as we know it demands – our morality is more complex – his objection is not a point against actual morality. What evidence is there that the Ross theory is mistaken? Morality permits telling harmless lies (even though there is a general obligation to speak the truth) when it is of personal benefit; the courts do not enforce "gratuitous" promises; morality permits pushing off a newcomer – thereby harming him – if I am clinging to a plank in a shipwreck; one is free to break one's marriage vow, thereby harming one's spouse, by getting a divorce, for no reason other than falling in love with someone else. The Ross conception of morality is thus too simple.

When is it rational to be interested in subscribing to some morality of the aversion/guilt/disapproval type? Answer: When, given we are fully informed, we *want* some such system, presumably both because of self-

## INTRODUCTORY COMMENTS

interest (for protection) and altruism. *Actual* moral systems often cannot make such an appeal. They may forbid harmless actions, like homosexuality or political liberalism. Reflections on the effects of a morality will make clear at which points an actual moral system needs change, and which aversions it is better to have stronger than certain others. If the moral system is of a kind we should most *want*, if fully informed, as compared with any other or none at all, it seems that its presence must be a benefit – contrary to one of Williams' claims.

Williams thinks that if an otherwise immoral (prohibited by the optimal code) action produces good, for example, contributes to art as Gauguin's actions did, then it is morally in the clear. Up to a point this is right, for an optimal moral system will presumably provide that some of its specific restrictions be set aside to avoid great harm or to produce great good. But how much harm/good? We can hardly be precise about this, but we must say "Quite a lot." Moreover, for this prospective good to suspend an otherwise stringent moral rule for a given person, the person must have good reason to think that his act is required to produce the good – not merely believe, or hope, that it will.

Williams argued that personal or aesthetic concerns also have a claim to be recognized, along with moral ones, and may be of greater weight. How should we arbitrate between them if there is a conflict? Suppose a person has very strong personal ambitions, stronger than his moral aversions and his aversion to the justified disapproval of others. Is there any reason we can offer why he should set a higher value on morality? I did not try to answer that question in this paper.

The fifth paper, Chapter 6, "Rationality, egoism, and morality" (1972), considers whether an optimal moral system – one we should want, if fully informed – might be an egoist one. A partial answer is that a person can enhance his well being more otherwise than by choosing an egoist system which requires that a person always act to maximize benefit to himself, irrespective of harm to others.

## B. NORMATIVE ETHICS: UTILITARIANISM

In the papers discussed above I argued that the kind of moral system that can be recommended to thoughtful people is one they will *want* for their society, provided their wants have been criticized by facing them vividly with relevant information. What kind of moral system would this be? In the next three papers I discuss forms of utilitarianism as proposals about the moral system to be chosen.

All types of utilitarianism agree that the rightness of an act is fixed by *some* relation to the production of *benefit*. A major difference between

## INTRODUCTORY COMMENTS

them is on the precise kind of relation to benefit rightness involves. The first two papers below are concerned with this controversy. Another disagreement is about what counts as a benefit: some hold that the only benefit is happiness, others that it is satisfaction of desire, still others that it is the occurrence of some good different from happiness but not necessarily desired. The third paper below is concerned with this dispute.

Of the two major theories about what relation the rightness of an act has to benefits, the first, usually called "act-utilitarianism," is very simple. It affirms that the to-be-preferred (most wanted) moral system requires an act, as being obligatory, if and only if performance of that act, of all the acts open to the agent, would produce the most good for everybody affected. (There are disagreements about whether it is actual good that is to be produced, or good that, on the agent's evidence, probably would be produced.) In the first paper below, I reject this theory partly on the ground it has intuitively unacceptable consequences (an argument I later rejected). The second theory, usually called "rule-utilitarianism," is more complex. It was first clearly stated, and distinguished from the "act" view, by Bishop Berkeley in *Passive Obedience*.

The first paper defending this view (Chapter 7) is "Some merits of one form of rule-utilitarianism" (1967). Its thesis is that we should regard an act as wrong if and only if it *would be* prohibited by a moral code concerned with various act-*types*, provided the general *recognition* of these prohibitions in the society *would* do most good. It is thus an "ideal" theory. (What does "general recognition" come to? Not universal *conformity*; such a code could be act-utilitarian. I suggested 90 percent acceptance, but I think we needn't worry too much about this.) The code must contain "rules" dealing with recurrent problem situations, be simple enough to be interiorized, not leave too much to discretion, not ignore the costs of teaching or conforming, or impose unbearable demands such as requiring one always to do what will produce the most good, like act-utilitarianism. (Most decisions in life, therefore, will not raise moral questions at all.) The system of rules presumably would require keeping contracts, making restitution for injuries, helping others in distress, and perhaps bringing about a just distribution of material goods. It will hardly be a worldwide universal code, but one that does best for a given society, with its institutions.

The second paper in this group (Chapter 8), "Fairness to indirect optimistic theories in ethics" (1988), was written twenty years later, and might be expected to show some progress! It is intended to add to, or amend, the foregoing theory, and to reply to various objections raised to the rule-utilitarian conception. It is pointed out that for a rational moral system to be prevalent in a society, someone must *teach* it and the

## INTRODUCTORY COMMENTS

teachers must have a view of the function of a moral code, and aim to teach proscriptions they think maximally beneficial to teach. We may assume that teachers will largely accept the moral code as it is in their society, and aim at only incremental changes. In the first paper I criticized act-utilitarianism partly on the ground that it is incompatible with our moral intuition; in the second I rely on more pragmatic criticisms, to the effect that if everyone were an act-utilitarian in his moral thinking it would be difficult to predict what other people might do – an obstacle to cooperation and to any planning based on expectations of the future behavior of others. Again, the problem of estimating utilities in a particular situation is such as to be a standing invitation to people to reach moral conclusions compatible with self-interest.

To objections that have been raised, I attempt to reply. It is said that it is impossible to make the utility comparisons necessary to apply the rule-utilitarian theory at a social level. A second objection is that it is irrational “rule-worship” to follow a moral rule if you happen to know which of two specific actions will have better results. Since the conception of a rule-utilitarian code does not identify what the rules will be, is there not the same difficulty in planning or cooperation that there is for the act-utilitarian? On the rule-utilitarian view, the specific rules to be followed are not decided by God, but necessarily by human agents. May they not disagree? Finally, it has been objected that it would hardly be useful to advocate following what *would* be the optimal code if everyone accepted it, when a large minority is bitterly opposed (to interracial marriage in South Africa, for example). I try to answer all these objections.

The third paper of this group (Chapter 9), “Two concepts of utility” (1982), is an attempt to adjudicate between two theories of what is intrinsically good: the happiness theory and the theory that it is desire-satisfaction in the sense of the occurrence of an event that one does (did or will) desire, irrespective of any enjoyment from it. The desire-satisfaction theory can add some conditions such as that the desire be fully informed but it is surely not clear that the satisfaction of desires is always a benefit to the agent. How about altruistic emotionally influenced desires? The concept of happiness is clear enough, and what makes one happy is often the same as what satisfies desire. Not always, however: A person may want honor or social status very strongly, when having them does not increase happiness. Moreover, I had argued in an earlier paper (“The Psychology of Benevolence and Its Implications for Philosophy,” *Journal of Philosophy*, 1976), on the basis of psychological theory, that what a benevolent/sympathetic person will want for others is *experiences* like those we want in ourselves, not just fulfillment of the others’ desires. I also argue that there is a larger problem for the desire theory,

## INTRODUCTORY COMMENTS

in that desires change as we mature. How, then, shall we decide which action best satisfies an agent's desires? The happiness theory has the merit of simplicity.

There is a further defense of the happiness theory in "Fairness to Happiness," *Social Theory and Practice*, 1989.

## C. UTILITARIANISM AND RIGHTS

Chapter 10, "The concept of a moral right and its function," aims to meet a widely agreed need to get a plausible definition of a "moral right." We can hardly get a useful dictionary definition, obviously synonymous with the definiendum, but must be satisfied with an "adequate replacement." I urge defining this term by employing the concept of moral obligation: "*X* has a moral right to *Y*" may be taken as "People in *X*'s society have a strong moral obligation to enable *X* to do, have, or enjoy *Y*, primarily because of the importance to people in *X*'s situation of being able to do, have, or enjoy things like *Y*." I proposed adding, "*X* morally ought to (or is permitted to) feel resentment if he is hurt or deprived because of the failure of others to discharge this obligation, and to feel unashamed to protest, and to take reasonable steps of protest, calculated to encourage others to have the motivation to enable anyone in a similar situation to do, have, or enjoy things like *Y*."

Since "ought" implies "can," it seems there can be no moral right to something that no strength of sense of moral obligation could bring about in a given society, for instance, to bring the standard of living in present-day India up to that of Sweden. What is one to say to this? We need to explain why there should be "rights-talk" at all. These queries need replies, which I try to offer.

The foregoing, however, is only a definition of the *concept* of a moral right. It does not yet tell us what are our rights, except to some extent by reference to the importance to some persons of having their rights recognized. Chapter 11, "Utilitarianism and moral rights" (1984), does discuss this issue. The above definition includes the notion of "moral obligation to . . ." The *rule-utilitarian* will say that one is obliged just in case an optimal moral code for the society would include such an obligation. It is quite consistent to be a rule-utilitarian about moral obligation and to adopt the above definition of a "moral right." We then come up with a normative rule-utilitarian theory about moral rights. One virtue of such a theory is that it enables us to identify our rights – for instance, the scope of the "right to life," and whether it is more binding than the right to free speech on political topics. I briefly criticize the views on moral rights of David Lyons and R.M. Hare.

## INTRODUCTORY COMMENTS

### D. DETERMINISM, EXCUSES, AND CHARACTER

I have argued that an act is right if a social *moral code* permitting that act would be best, do most good, and I have said that a person has a moral code forbidding a kind of action if and only if he has an aversion to doing that sort of thing, just in itself, *and* has a disposition to feel guilty if he does it, and to disapprove of others who do, and thinks these attitudes are justified in some appropriate sense. We have to recognize some qualifications in the clauses about guilt feelings and disapproval, however, as I explain in Chapter 12, "A utilitarian theory of excuses" (1969), just as, in the law, an unlawful act is not liable to punishment unless there was *mens rea*, so that an act is legally excused if it occurred owing to mistake of fact or accident. The rule-utilitarian will consider these qualifications, because which ones we allow will make a difference to the desirability of a certain moral system. A moral system works in part because people do not like to be disapproved by others, and because people do not like to feel guilty, or take a dim view of themselves as persons. Thus guilt feelings and condemnation by others motivate people to do what is socially beneficial. But marshaling these condemning attitudes will not tend to prevent certain types of behavior: injuring another person accidentally, or failing to do something impossible. These facts lead us to recognize a general principle: A person should be morally excused for some untoward act if it does not manifest some *defect of character* – at least if we take this to imply a defect in motivation. If a person's aversions to acting in certain ways are already adequate (there is no defect of character), the moral system has done its job. If this is not the case, then guilt feelings and criticism by others can help improve the level of motivation. We should note that criticism can help even if all actions are caused; this theory of excuses is not inconsistent with determinism.

The same position should be taken by the criminal law, as I argue in the twelfth paper, Chapter 13, "A motivational theory of excuses in the criminal law" (1985).

According to the criminal law, a person is liable to punishment if he performs a forbidden act (*actus reus*), unless it is *justified* (perhaps because greater social good is done than by omitting it), or is *excused* (say, by ignorance of fact or by accident), or the agent is not responsible (insane or an infant). The thesis of this paper is that an act must be excused (even if an unjustified breach of law) if it fails to show a defect of motivation. My thesis is thus two-fold: for the most part excuses recognized by the law are in fact of this kind, and such excuses are what it *ought* to recognize.

Older legal theory said an act is punishable only if it springs from an



## INTRODUCTORY COMMENTS

“evil will,” but present law seems to say an act is punishable if the agent *knowingly* or *recklessly* does a legally forbidden thing. It could as well say that it is punishable if the act shows defect of character (motivation) – *indifference* to harm, or at least to the kinds of acts forbidden, or to the fact that they are forbidden. If we follow this line, does it mean the judge must speculate about motivation? If the judge must make the decisions the Model Penal Code says he must now make, he is already speculating! In fact, we draw inferences about defective motivation from overt behavior all the time. The motivational theory implies that the following excuses should be recognized: accidents; mistake of fact; mistake of law (although here most judicial opinion clearly does not agree); involuntary intoxication; duress; provocation (as a mitigating excuse); the law of attempts; some kinds of insanity. All these permit a no-defect-of-motivation explanation.

Utilitarians will hold that the main benefit of the legal system is to reduce harmful behavior. It does this primarily by making the standards known and by deterrence (putting threats of punishment on the side of lawful behavior). If a person has broken the law but not because of defect of motivation, no benefit is gained from punishing him; there is no deterrence. His motivation may already be where it ought to be. Some people think that we should not forget deterrence of *others* whose motivations may not be perfect, and that knowledge that there is a system of excuses (that some unlawful acts are not punished) reduces respect for the law; but this remains to be shown, and a system of strict liability would make life intolerable.

The preceding articles try to show that the concept of a trait of character is important both for normative ethics and for legal philosophy. We therefore need a conception of a trait of character consistent with the reasoning about both moral and legal excuses, and fitting into common sense conceptions and contemporary psychological theory. “Traits of character: A conceptual analysis,” Chapter 14, is intended to provide this. This paper argues that we must consider traits of character as dispositions, in contrast with the view of some that trait names (“sympathetic,” “courageous”) merely summarize past behavior, or with the view of others that they are dispositions only to manifest specific types of behavior. In fact, it is clear that traits of character like sympathy, fidelity, and truthfulness are *motivational*, wants or aversions, and enter into explanations of intentional behavior just as do other wants and aversions. We infer or deny traits on the basis of a single example of behavior, which we could not do on the specific behaviors theory.

If we define a moral “virtue” as a trait of character beneficial to society, the question arises whether *all* moral virtues are motivational, as the